DOCUMENT RESUME

ED 335 662 CS 010 687

AUTHOR VanLeirsburg, Peggy
TITLE Socio-Educational Influences on Standardized Reading Tests, 1900-1991. Literacy Research Report No. 11.
INSTITUTION Northern Illinois Univ., DeKalb. Curriculum and Instruction Reading Clinic.
PUB DATE Nov 91
NOTE 14p.; Based on a dissertation "The Historical Development of Standardized Reading Tests in the United States, 1900-1991," completed at Northern Illinois University.
AVAILABLE FROM Northern Illinois University, The Reading Clinic, 119 Graham, DeKalb, IL 60115 ($3.50).
PUB TYPE Reports - Research/Technical (143) -- Historical Materials (060)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Educational History; *Educational Testing; Elementary Secondary Education; *Reading Tests; *Social Influences; *Standardized Tests; Testing Problems
IDENTIFIERS Historical Background

ABSTRACT

Since 1915, the movement to test reading ability has grown in importance and pervasiveness. Testing is relied on to offer quantitative information, to determine the usefulness of instructional approaches, and to help ensure accountability. The demand for tests arose during the period from 1990 to 1992 when school became compulsory and educators faced new challenges with the range of abilities and diversity of cultures that individuals brought to the public school system. In 1915, W. S. Gray published the first standardized reading test. Later, other reading tests were published to measure comprehension. During the period from 1930 to 1965, schools relied heavily on standardized tests to diagnose disability, predict success, place students in special programs, and demonstrate instructional effectiveness. From 1966 to 1991, federal legislation and judicial decisions requiring public schools to provide free educational services contributed to the further pervasiveness of testing. Although reading tests have not changed significantly in the past 70 years, there have been advances in statistical sophistication relative to reliability and validity. But there is a growing dichotomy between recent advances in understanding the reading process and the tests used to measure reading achievement. The following five recommendations can help to improve large-scale reading assessment: (1) tests should be administered with a specific purpose; (2) the amount of testing should be reduced; (3) knowledge about tests should be improved and increased; (4) standardized tests should be carefully chosen; and (5) assessment information should be used for instructional purposes at a personal level. (One table of data is included and 32 references are attached.) (PRA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED335662

Northern Illinois University

Literacy Research Report No. 11

SOCIO-EDUCATIONAL INFLUENCES ON STANDARDIZED READING TESTS, 1900-1991

Peggy VanLeirsburg
Elgin Public Schools

November 1991

Literacy Research and Reports

Curriculum and Instruction
Reading Clinic—119 Graham
DeKalb, Illinois 60115

CS010687

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Peggy VanLeirsburg

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

Literacy Research Report No. 11

SOCIO-EDUCATIONAL INFLUENCES ON STANDARDIZED READING TESTS, 1900-1991

Peggy VanLeirsburg
Elgin Public Schools

November 1991

Northern Illinois University
Reading Clinic--119 Graham Hall
DeKalb, Illinois 60115
815-753-1416

This report is based on the author's dissertation, The Historical Development of Standardized Reading Tests in the United States, 1900-1991, completed at Northern Illinois University in 1991.

Socio-Educational Influences on

Standardized Reading Tests, 1900-1991

Since its birth in 1915, the movement to test reading ability has grown in importance and pervasiveness. Testing has become a business that is relied on to offer quantitative information, to determine usefulness of instructional approaches, and to help ensure accountability. Rough estimates suggest that a large percentage of children in this country receive more than 2,000 test items every year--including items on standardized tests, basal tests, and teacher-made tests (Tierney, Carter, & Desai, 1991). A typical student may spend two to six hours each year taking tests throughout his or her elementary and high school years; special education students may experience two to three times more testing. Tests which accompany textbooks, local instructional tests, and teacher-made tests are not included in this estimate (Wigdor & Garner, 1982).

Most school systems use standardized tests. National goals by the year 2000 include: preparedness to learn upon school entry; increased competency at the end of grades 4, 8, and 12; 90% high school graduation rate; and adult literacy. Documenting achievement of cognitive goals seems to rely on standardized testing to "prove" they have been reached. Standardized testing has become entrenched in our schools and society. Reading is one of the most frequently measured academic abilities (Smith, 1986). The publication of standardized reading tests and achievement test batteries with reading subtests tests by decade, not including test revisions, is reported in Table 1. These data were compiled from the following sources: Volumes 1 through 10 of The Mental Measurements Yearbooks (Buros, 1938, 1940, 1949, 1953, 1959, 1965, 1972, 1978; Conoley & Kramer, 1989; Mitchell, 1985), Volumes 1 through 3 of Tests in Print (Buros, 1961, 1974; Mitchell, 1983), and Volumes 1 and 2 of Reading Tests and Reviews (Buros, 1968, 1975).

Table 1

Standardized Reading Tests and Achievement Test Batteries
For Elementary Students Published in the United States by Decade

Decade	Number	Cumulative Number	# Out Of Print	Cumulative # Out Of Print
1910-1919	5	5	3	3
1920-1929	21	26	10	13
1930-1939	33	59	15	28
1940-1949	12	71	2	30
1950-1959	21	92	4	34
1960-1969	28	120	0	34
1970-1979	20	140	0	34
1980-1989	8	148	0	34

The Past

Assessment in reading is not a value-neutral enterprise; it reflects the value commitments and ideology apparent in societal and cultural standards of the time (Archbald & Newmann, 1988). Historical events provide a frame of reference for assessment trends as researchers and practitioners respond to the critical socio-educational demands of our society.

1900-1929

The demand for tests arose during a period when school was made compulsory (Glaser, 1981). The population of public schools rose from 18 million in 1910 to around 25 million in 1930 as a result of mandatory attendance laws, the incorporation of kindergarten, and child labor laws (Rippa, 1988). Educators faced new challenges with the range of abilities and diversity of cultures that individuals brought to the public school system. Testing to predict which students would benefit from specialized instruction, vocational programming, or college preparation was seen as part of the solution (Glaser, 1981).

instruction, vocational programming, or college preparation was seen as part of the solution (Glaser, 1981).

Reading was the last academic subject to yield to the testing movement (Smith, 1986). In 1915, W. S. Gray, a student of Thorndike, published the first standardized reading test, the Gray Standardized Oral Reading Paragraphs. Gray's work showed that a test of oral reading was cumbersome and inefficient to administer and score in a group setting. Further, reading was a complex task to measure. Following the work of Gray, other reading tests (e.g., Brown, 1914; Kelly, 1916; Courtis, 1917; and Monroe, 1918) were published that required silent reading on the part of the student. The original group reading tests showed commonalities. They required the student to read silently, supply a short answer, make one choice from supplied multiple choices, or retell a story to measure comprehension. Tests were usually timed; speed of reading was a desired goal.

1930-1965

Three main social and cultural forces influenced standardized reading testing from 1930 to 1965. The first was the Depression and the collapse of the stock market in 1929. World War II and the ensuing effort to establish the United States as an international leader was a second social and cultural force. Federal funds were needed to support the war effort and were not channeled into education. The third major social and cultural factor affecting the standardized testing of reading was the rapid growth of communications and technology. The development of the IBM 805 scanner in 1935 fixed the place of objective, multiple-choice tests by reducing the cost of scoring to one-tenth that of hand-scoring (Pearson & Dunning, 1985).

Criticism of standardized reading tests continued during the time period 1930 to 1965. However, reliance on these tests for the measurement of reading skills became widely used in schools. Tests were used to diagnose disability, predict success, place students in special programs, and demonstrate instructional effectiveness.

1966-1991

Standardized testing became firmly entrenched in the system of the United States from 1966 to 1991. Federal legislation and judicial decisions contributed to the further pervasiveness of testing. Equality of opportunity for individuals in a diverse society was the ideal supported by legal action. In his 1964 State of the Union address, President Lyndon B. Johnson declared war on poverty and challenged education to play a major role (Cremin, 1988). Congress responded by passing an unprecedented record of legislation for education (Rippa, 1988).

Judicial decisions had a direct impact on the schools rendering it illegal for schools to deny an education to any student. *Pennsylvania Association for Retarded Children (PARC) v. Commonwealth of Pennsylvania* (1971) and *Mills v. Board of Education* (1972) set a precedent by requiring school districts to provide an appropriate education to children with physical, mental, and emotional handicaps, no matter how severe. The PARC and Mills decisions were reaffirmed in 1975 when Public Law 94-142 was signed by President Gerald Ford. Public school systems were required to provide free educational services and enormous amounts of standardized testing to qualify or disqualify exceptional children ages 3 to 21 in the least restrictive environment. In the 1976-77 school year, about 3.7 million students were served in federally supported special education programs; by 1987-88, the total was over 4.4 million students, or over 11% of the school population (Snyder, 1989).

Federal influence demanded data and accountability in the form of standardized test scores. The use and pervasiveness of testing in the educational setting and in society increased. Standardized tests of reading ability have increased in number, but have not changed in basic format and theoretical basis. Test authors attempted to compile and measure a myriad of reading skills (Roser, 1984).

Summary

The evidence suggests that reading tests have not changed significantly in the past 70 years. There have been advances in statistical sophistication relative to reliability and validity. Farr and Carey (1986) credit the invention, development, and extensive use of test scoring machines and computer

summaries of test information with the radical change in speed with which test results are determined. However, with the exception of changes in some readiness tests to reflect current knowledge about the literacy development of pre-school children, tests themselves look similar to those developed during the 1920s.

The Present

Tests are administered to millions of elementary children each year and reading is probably the most frequently tested behavior (Farr & Carey, 1986). The importance of reading is reflected by the prevalence of testing of reading. National progress examinations, state competency tests, achievement batteries, and special education program requirements include a test of reading. Many schools that give standardized achievement batteries administer only the reading and mathematics tests. Programs that include standardized reading testing are firmly entrenched in our schools and society. Yet, there is a growing dichotomy between recent advances in understanding the reading process and the tests used to measure reading achievement.

The Future

The keys to meeting the reading assessment needs of a diverse student population are a flexible approach to assessment and an improved knowledge base on the part of educators and the public (Garcia & Pearson, 1990). Costa (1989) confirmed the tenets of flexible assessment and improved knowledge base. He further outlined three requirements for the redesigning of assessment: re-establish the school as the locus of accountability; expand the range and variety of assessment techniques; and systematize the variety of assessment procedures.

However, the socio-economic events that have developed in the United States since the early 1900s influenced a standardized testing legacy that will not easily be changed. It is unlikely that test publishers, educators, or the public will be satisfied with assessment measures that report anything but normed data in a single score. It is unlikely that the convenience of these tests can be matched; standardized tests are economical and can be administered to large groups in a relatively short time and

scored quickly by computers that offer a wide variety of data profiles.

The promise of improved large-scale assessment of reading has begun in some states. These changes are still under development and not generally accepted or understood by eductors or the public. Hopefully, this innovative trend will continue and surpass the use of typical group standardized reading tests. While this gains impetus, the following recommendations are suggested for large-scale reading assessment.

1. Testing should not be administered without a specific purpose. Some districts administer the same achievement battery each spring because it has become tradition, not necessity. Instead, testing reflective of local curricula and diversity that will produce useful data for a specific reason is encouraged.

2. The amount of standardized testing of reading should be reduced. If one measure offers adequate information, a second test need not be given.

3. Knowledge about standardized tests should be improved and increased. Assessment literacy is encouraged at all levels. School districts and teacher-training programs need to become aware of the fact that at all levels, from principals to parents, many individuals pretend to be knowledgeable about statistics that actually confuse them.

4. Standardized tests should be carefully chosen. The tests should reflect current knowledge about reading and reflect the curriculum in a given geographical area.

5. While some amount of testing may be necessary to report statistics for federal and state use, the most useful assessment information for instructional purposes occurs at a more personal level involving an individual or classroom.

References

Archbald, D. A., & Newmann, F. M. (1988). Beyond standardized testing. Reston, VA: National Association of Secondary School Principals.

Brown, H. A. (1914). The measurement of efficiency of instruction in reading. Elementary School Teacher, 14, 477-490.

Buros, O. K. (Ed.). (1938). The nineteen thirty-eight mental measurements yearbook. Highland Park, NJ: Gryphon Press.

Buros, O. K. (Ed.). (1940). The nineteen forty mental measurements yearbook. Highland Park, NJ: Gryphon Press.

Buros, O. K. (Ed.). (1949). The third mental measurements yearbook. Highland Park, NJ: Gryphon Press.

Buros, O. K. (Ed.). (1953). The fourth mental measurements yearbook. Highland Park, NJ: Gryphon Press.

Buros, O. K. (Ed.). (1959). The fifth mental measurements yearbook. Highland Park, NJ: Gryphon Press.

Buros, O. K. (Ed.). (1961). Tests in print. Highland Park, NJ: Gryphon Press.

Buros, O. K. (Ed.). (1965). The sixth mental measurements yearbook. Highland Park, NJ: Gryphon Press.

Buros, O. K. (Ed.). (1968). Reading tests and reviews. Highland Park, NJ: Gryphon Press.

Buros, O. K. (Ed.). (1972). The seventh mental measurements yearbook. Highland Park, NJ: Gryphon Press.

Buros, O. K. (Ed.). (1974). Tests in print II. Highland Park, NJ: Gryphon Press.

Buros, O. K. (Ed.). (1975). Reading tests and reviews II. Highland Park, NJ: Gryphon Press.

Conoley, J. C., & Kramer, J. J. (Eds.). (1989). The tenth mental measurements yearbook. Lincoln: University of Nebraska Press.

Costa, A. L. (1989). Re-assessing assessment. Educational Leadership, 46, 2.

Courtis, S. A. (1917). The problems of measuring ability in silent reading. American School Board Journal, 54, 17-18, 81.

Cremin, L. A. (1964). American education: The metropolitan experience. New York: Harper & Row.

Farr, R., & Carey, R. F. (1986). Reading: What can be measured? (2nd ed.). Newark, DE: International Reading Association.

Garcia, G. E., & Pearson, P. D. (1990). The role of assessment in a diverse society. Unpublished manuscript.

Glaser, R. (1981). The future of testing. American Psychologist, 36, 923-936.

Gray, W. S. (1915). Standard oral reading paragraphs. Indianapolis, IN: Bobbs-Merrill.

Kelly, F. J. (1916). The Kansas silent reading tests. Journal of Educational Psychology, 7, 63-80.

Mitchell, J. V. (Ed.). (1983). Tests in print III. Lincoln: University of Nebraska Press.

Mitchell, J. V. (Ed.). (1985). The ninth mental measurements yearbook. Lincoln: University of Nebraska Press.

Monroe, W. S. (1918). Monroe's standarized silent reading tests. Journal of Educational Psychology, 9, 303-312.

Pearson, P. D., & Dunning D. (1985). The impact of assessment on reading instruction. Illinois Reading Council Journal, 13, 19-29.

Rippa, S. A. (1988). Education in a free society. New York: Longman.

Roser, N. L. (1984). Teaching and testing reading comprehension. In J. Flood (Ed.), Promoting reading comprehension (pp. 48-60). Newark, DE: International Reading Association.

Smith, N. B. (1986). American reading instruction. Newark, DE: International Reading Association.

Snyder, T. D. (1989). Digest of educational statistics. Washington, DC: U.S. Government Printing Office.

Tierney, R. J., Carter, M. A., & Desai, L. E. (1991). Portfolio assessment in the classroom. Norwood, MA: Christopher-Gordon.

Wigdor, A. K., & Garner, W. R. (1982). Ability Testing: Uses, consequences, and controversies, Part II, Documentation section. Washington, DC: National Academy Press.

LITERACY RESEARCH AND REPORTS ORDER FORM

Northern Illinois University
The Reading Clinic - 119 Graham
DeKalb, IL 60115
815-753-1416

No. of Copies

No. 1 Portfolio Assessment: A Survey Among Professionals
by Jerry L. Johns & Peggy VanLeirsburg _____

No. 2 Adults' Perceptions of Change in Reading Ability Across the Life Span
by M. Cecil Smith & Wesley C. Covalt _____

No. 3 Teacher Behaviors During Reading Instruction in Australian and American Classrooms
by Roberta L. Berglund, James P. Raffini, & Lorraine McDonald _____

No. 4 Perceptions of Preservice and Inservice Teachers Regarding Test-Taking Procedures and Test-Wiseness Programs
by Jerry L. Johns & Susan J. Davis _____

No. 5 Disappearance of Drawings in Children's Writing: A Natural Development or a Natural Disaster
by Gail Neu & Roberta L. Berglund _____

No. 6 An Analysis of the History of Writing and the Development of Children's Literacy Acquisition
by June E. Barnhart _____

No. 7 Developing a Love of Reading: What Helps, What Hurts
by Roberta L. Berglund, Richard J. Telfer, & Joan E. Heimlich _____

No. 8 Three Reading Groups: An American Educational Tradition
by Susan J. Davis _____

Each report costs $3.50 (postage included). Actual postage costs will be added for large orders or overseas shipments. Please remit in U.S. funds. Please make check or money order payable to NIU Reading Clinic.

Ship to:

Name _____

Address _____

_____

Amount Enclosed

City _____ State _____ Zip _____ $_____

Country _____

☐ Place me on your mailing list.

(more reports on next page)

No. of Copies

No. 9	Annotated Bibliography of Whole Language Resources by Amy K. Aufmann & Mary Ann Wham	_____
No. 10	How Professionals View Portfolio Assessment by Jerry L. Johns & Peggy VanLeirsburg	_____
No. 11	Socio-Educational Influences on Standardized Reading Tests, 1900-1991 by Peggy VanLeirsburg	_____
No. 12	Assessment Literacy: Perceptions of Preservice and Inservice Teachers Regarding Ethical Test-Taking Procedures by Peggy VanLeirsburg & Jerry Johns	_____